A Wavelet-Based Reduced Rank Kalman Filter for Assimilation of Stratospheric Chemical Tracer Observations

LUDOVIC AUGER[*]

Météo-France, Toulouse, France

ANDREW V. TANGBORN

Global Modeling and Assimilation Office, NASA GSFC, Greenbelt, and JCET, University of Maryland, Baltimore County, Baltimore, Maryland

ABSTRACT

A suboptimal Kalman filter system that evolves error covariances in terms of a truncated set of wavelet coefficients has been developed for the assimilation of chemical tracer observations of CH_4. The truncation is carried out in such a way that the resolution of the error covariance is reduced only in the zonal direction, where gradients are smaller. Assimilation experiments, which lasted 24 days and used different degrees of truncation, were carried out. These experiments reduced the number of elements in the covariance matrix by 90%, 97%, and 99% and the computational cost of covariance propagation by 80%, 93%, and 96%, respectively. The difference in both error covariance and the tracer field between the truncated and full systems over this period was not found to be growing after about 5 days of assimilation. The largest errors in the tracer fields were found to occur in regions of largest zonal gradients at times when observations were made in the immediate vicinity.

1. Introduction

The development of suboptimal Kalman filter schemes for data assimilation has been motivated by the need to reduce the computational expense of evolving error covariances. These techniques can be roughly divided into representation by direct reduced rank expansion, Monte Carlo, and parameterized flow-dependent covariance methods. Truncated expansion methods use an efficient representation of the error covariance and its propagator, such as singular vectors and eigenvectors of the system matrix (Tippett et al. 2000; Farrell and Ioannou 2001; Cohn and Todling 1996; Ehrendorfer and Tribbia 1997), coarse-grid approximation (Fukumori and Malanotte-Rizzoli 1995), EOFs (Cane et al. 1996), or wavelets (Chin et al. 1999; Tangborn and Zhang 2000). These can then be truncated to a few leading terms, resulting in a low-dimensional system. The Monte Carlo or ensemble Kalman filter method (Evensen 1994; Houtekamer and Mitchell 1998) avoids the ex-

pense of propagating error covariances by estimating them from an ensemble of short-range forecasts. This approach also has the advantage of obtaining current ensemble statistics instead of relying on a tangent linear model (TLM) to evolve error statistics. However, computational limitations generally allow for ensemble sizes in the hundreds, and therefore the possibility of spurious solutions require the truncation of covariances for large distances. Parameterized flow-dependent error covariance (Riishøjgaard 1998) assumes that the background error correlation has essentially the same shape as the background field. This approach has the potential to estimate anisotropic correlations with only a slightly higher computational cost than isotropic error correlations.

All of these approaches have the goal of estimating local, anisotropic error correlations. An important step developing an error covariance estimation scheme is determining whether a given method has the potential to be implemented in a realistic large-scale assimilation system. A wavelet basis is a logical means for representing both the error covariances and the TLM because the approximation is local in position and scale. Not only is this representation highly compressive for anisotropic data, but it gives explicit information on the location and scale of the most important covariance structures. This allows for adaptive schemes to truncate the wavelet expansion to a small number of significant

---

* Current affiliation: Laboratoire de Meteorologie Dynamique, Ecole Normale Superieure, Paris, France.

Corresponding author address: Dr. Andrew V. Tangborn, Global Modeling and Assimilation Office, Code 900.3, Goddard Space Flight Center, Greenbelt, MD 20771.
E-mail: tangborn@gma.gsfc.nasa.gov

terms and potentially could provide insight into the relationship between important features in the error covariance and background fields. Further, wavelet transforms apear to be an appropriate approach for solving this sort of evolution equation given their successful application to computational fluid dynamics during the past 10 yr (Farge 1992).

Recent work on Kalman filtering of chemical tracer observations (Ménard et al. 2000; Ménard and Chang 2000) has provided a benchmark with which to test suboptimal schemes for tracer assimilation. In this system, chemical species observations (e.g., CH_4) from limb-sounding instruments [Cryogenic Limb Array Etalon Spectrometer (CLAES) and Halogen Occultation Experiment (HALOE)] are assimilated into a global two-dimensional transport model on isentropic surfaces. All of the necessary specifications, including initial, model, and observational error covariances, are designed and tested in this work. Testing of a number of simplifying assumptions that eliminate vertical error correlations of observations, diabatic effects, vertical mixing, and chemical sources and sinks is also done. The results of this work show that the error covariances do have some state dependence. Ménard and Chang (2000) also discovered that there is significant loss of error variance due to ''spurious dissipation in the small-scale covariance structures.'' This problem was overcome by introducing a variance correction term that keeps the total error variance constant during the forecast propagation step (and before the addition of model error). The complexities of propagating the error covariance in wavelet space make this variance correction step unviable. Therefore the variance loss issue is not considered in the present work.

In this study, we have taken the Kalman filter system developed by Ménard et al. (2000) and projected the error covariances and covariance propagator (transport model) onto an orthonormal wavelet basis. The resulting system then carries out the propagation of error covariances in ''wavelet space.'' Several different levels of truncation are used, and all of the results are compared with the *benchmark* results in terms of error covariance, the constituent field, and computational time. While the full system does not represent the true state, its behavior is well understood, and we seek to understand how this low-dimensional representation affects the covariance evolution and how this in turn affects the assimilated constituent field.

2. The transport advection scheme and constituent assimilation system

The details of the Kalman filter assimilation scheme can be found in Ménard et al. (2000), and we give a short summary here. The evolution of a chemical constituent field along an isentropic surface is governed by

$$\frac{\partial \boldsymbol{\mu}^t}{\partial t} + \mathbf{V} \cdot \boldsymbol{\nabla} \boldsymbol{\mu}^t = f, \tag{1}$$

where $\boldsymbol{\mu}^t$ is the true tracer mixing ratio, $\mathbf{V}$ is the horizontal wind field along an isentropic surface, and $\boldsymbol{\nabla}$ is the horizontal gradient operator restricted to the isentropic surface. The true state of any constituent field will depend on chemical reaction (source and sink) and diffusion and is represented by the random forcing term f. The model used to make forecasts is conservative and ignores the sources, sinks, and diffusion, so that f is set to zero in the transport code.

The numerical model used to solve the transport equation is adapted from the Lin and Rood (1996) scheme, which uses the conservative form

$$\frac{\partial \boldsymbol{\mu}}{\partial t} + \boldsymbol{\nabla} \cdot (\boldsymbol{\mu}\mathbf{V}) = \boldsymbol{\mu}\boldsymbol{\nabla} \cdot \mathbf{V} \tag{2}$$

and is solved by an operator splitting approach. The system is solved offline using wind analyses $\mathbf{V}$ produced by the Goddard Earth Observing Data Assimilation System (GEOS-DAS). The discretized equations have the form

$$\boldsymbol{\mu}_{k+1} = \mathbf{M}_k \boldsymbol{\mu}_k + \varepsilon_k^q, \tag{3}$$

where $\boldsymbol{\mu}_k$ is the mixing ratio at time t_k, and $\mathbf{M}_k$ the discretized transport model from t_k to t_{k+1}. The model error ε_k^q in (3) comes from discretization and the neglected right-hand side of 1; ε_k^q is assumed to have zero bias and be normally distributed, with associated covariance $\mathbf{Q}_k$. Ménard et al. (2000) showed that state-dependent model error is a better choice than an isotropic form. There, the model error covariance is assumed to have the form

$$\mathbf{Q}_k(i, j) = \delta^2 \boldsymbol{\mu}_k^a(i)\boldsymbol{\mu}_k^a(j)\mathbf{C}^q(i, j), \tag{4}$$

where the correlation $\mathbf{C}^q(i, j)$ is given by

$$\mathbf{C}^q(i, j) = \exp\left(\frac{-|\mathbf{r}_i - \mathbf{r}_j|}{L}\right), \tag{5}$$

where L is the correlation length scale, $\mathbf{r}_i$ and $\mathbf{r}_j$ are position vectors of two points on the sphere, and δ is a relative error parameter. The initial error covariance matrix takes the same form as the model error covariance

$$\mathbf{P}_0(i, j) = \gamma^2 \boldsymbol{\mu}_0(i)\boldsymbol{\mu}_0(j)\mathbf{C}(i, j), \tag{6}$$

where $\mathbf{C}(i, j)$ is the same as in (5).

The observations of mixing ratios at time t_k, $\boldsymbol{\mu}_k^o$ can be related to the true mixing ratio by

$$\boldsymbol{\mu}_k^o = \mathbf{H}_k \boldsymbol{\mu}_k^t + \varepsilon_k^o, \tag{7}$$

where ε_k^o is the observation error, which includes contributions from measurement, retrieval, and representativeness errors. The observation error is assumed unbiased, normally distributed, and state independent and has associated error covariance matrix $\mathbf{R}_k^o$.
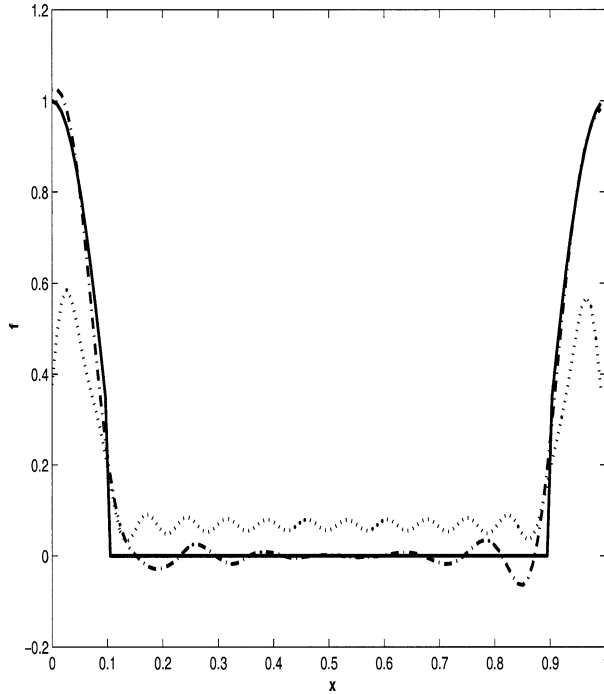
The Kalman filter algorithm consists of the forecast step, which updates the constituent forecast, $\boldsymbol{\mu}^f$, and its associated forecast error covariance, $\mathbf{P}^f$:

$$\boldsymbol{\mu}_{k+1}^f = \mathbf{M}_k \boldsymbol{\mu}_k^a, \tag{8}$$

$$\mathbf{P}_{k+1}^f = \mathbf{M}_k \mathbf{P}_k^a \mathbf{M}_k^T + \mathbf{Q}_k, \tag{9}$$

and in the analysis step, the constituent analysis, $\boldsymbol{\mu}^a$, and the analysis error covariance, $\mathbf{P}^a$ are calculated

$$\boldsymbol{\mu}_{k+1}^a = \boldsymbol{\mu}_{k+1}^f + \mathbf{K}_{k+1}(\boldsymbol{\mu}_{k+1}^o - \mathbf{H}_{k+1}\boldsymbol{\mu}_{k+1}^f), \tag{10}$$

$$\mathbf{P}_{k+1}^a = (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1})[(\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1})\mathbf{P}_{k+1}^f]^T$$
$$+ \mathbf{K}_{k+1}\mathbf{R}_{k+1}^o \mathbf{K}_{k+1}^T, \tag{11}$$

where the Kalman gain matrix $\mathbf{K}_k$ is given by

$$\mathbf{K}_{k+1} = (\mathbf{H}_{k+1}\mathbf{P}_{k+1}^f)^T(\mathbf{H}_{k+1}\mathbf{P}_{k+1}^f \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1}^o)^{-1}. \tag{12}$$

The error covariance dynamics [Eq. (9)] are computed by Ménard et al. (2000) using a two-step scheme that corrects for variance loss. The forecast error covariance is initially computed as

$$\mathbf{P}_{k+1}^f = \mathbf{M}_k \mathbf{P}_k^a \mathbf{M}_k^T, \tag{13}$$

and the error variance is treated as a conserved scaler and computed as

$$\mathbf{V}_{k+1}^f = \mathbf{M}_k \mathbf{V}_k^a. \tag{14}$$

The computational grid for this assimilation is 72×46 (for the zonal and meridional directions).

3. Wavelet representation of error covariances

Wavelets are a family of basis functions with local support. They are generated from scaling functions that satisfy the recursion relationship

$$\phi(x) = \sum_k c_k \phi(2x - k), \tag{15}$$

where the choice of filter coefficients c_k determines the properties of the resulting scaling functions and wavelets. A small number of nonzero c_k will give rise to more localized wavelets, and a larger number of c_k will produce smoother (i.e., continuous higher derivatives) wavelets. The basic wavelet function, ψ, for a given set of filter coefficients can be derived from scaling functions by taking differences:

$$\psi(x) = \sum_k (-1)^k c_{1-k} \phi(2x - k). \tag{16}$$

The representations of different scales and locations are then obtained by dilating and translating $\psi(x)$:

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \tag{17}$$

where j refers to dilation (scale), and k refers to translation (location). The form of Eq. (17) implies that the resolution scale decreases by a factor of 2 for each increment in j.

We are employing the discrete wavelet transform be-cause we wish to represent a covariance matrix by its wavelet coefficients. The pyramidal wavelet transform of Mallat (1989) recursively determines the coefficients at each scale, from finest to coarsest, by calculating weighted averages and differences at each scale. The transform can be represented by a matrix, $\mathbf{W}$, though the matrix itself is never actually formed. We write the transform in matrix form as

$$\hat{\mathbf{f}} = \mathbf{W}\mathbf{f}, \tag{18}$$

where $\mathbf{f}$ is a vector in physical space, and $\hat{\mathbf{f}}$ are the corresponding wavelet coefficients.

The discrete wavelet transform requires vectors with lengths that are powers of 2, and NWP models generally do not conform to this restriction. We therefore employ a constant value padding that extends the vector length from N to N_p using the value at the last entry in $\mathbf{f}$. This method is demonstrated on a one-dimensional example. We start with the function on [0, 1]:

$$f(x) = \begin{cases} \cos(2\pi x) & \text{for } x \leq 0.1 \text{ or } x \geq 0.9 \\ 0 & \text{elsewhere,} \end{cases}$$

which results in a localized periodic function that crosses the periodic boundary. The function is discretized using a grid of 115 equally spaced points, $f_i = f(x_i)$, then extended to 128 points using the value at f_{115} to pad the extended array so as to form the vector $\mathbf{f}_p$. The discrete wavelet transform is then applied:

$$\hat{\mathbf{f}} = \mathbf{W}\mathbf{f}.$$

The vector $\hat{\mathbf{f}}_p$ is then truncated to 16 terms (a more severe truncation than we are using in the assimilations) and transformed back to physical space using the inverse discrete wavelet transform. Figure 1 shows the original $\mathbf{f}$ and the 16-term reconstruction, for both wavelet and Fourier transforms. While there are Gibbs-type oscillations near the most abrupt part of the function, there is no loss of information accross the periodic boundary as a result of the padding. This contrasts to the more global errors that result from the 16-term Fourier reconstruction.

Representation of a function in two dimensions is similar to Fourier or other bases; we use the product of two functions:

$$P(x, y) = \sum_{j_k, k_x} \sum_{j_y, k_y} \hat{P}_{j_x, k_x, j_y, k_y} \psi_{j_x, k_x}(x)\psi_{j_y, k_y}(y), \tag{19}$$

where $\hat{P}_{j_x, k_x, j_y, k_y}$ is the wavelet coefficient for scales j_x, j_y and locations k_x, k_y. The scales and translations in each direction are varied independently of each other.

In order to understand the structure of the matrix that is to be represented by a wavelet basis, we plot an example of the entire error covariance in Fig. 2 after a 9-day assimilation. The (x, y) coordinates are collapsed into a single vector (with the zonal direction incremented in the inner loop). The indexing for the constituent field is then
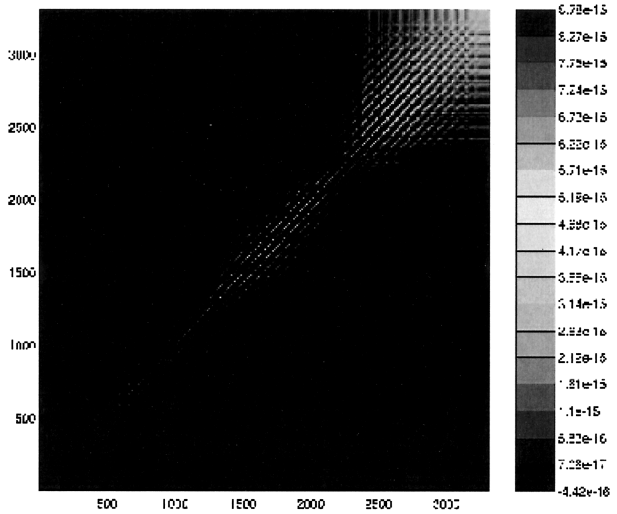
FIG. 2. Error covariance for the entire two-dimensional domain after a 9-day assimilation. The two-dimensional (72 × 46) gridded field represented by this matrix is collapsed into a single vector of length 3312. The error variances are represented by the main diagonal, with the most southerly points at diagonals 1–72 and the most northerly grid points at diagonal elements 3241–3312. The northern grid error variances are significantly larger than those in the southern region because the CLAES latitudinal coverage is 80°S–34°N during the period of the assimilation. The lack of observations in the high latitudes results in a larger error variance there.

FIG. 1. An example of the effect of padding on the reconstruction of a function from truncated wavelet and Fourier expansions. The periodic function $f(x)$ is discretized on a grid of 114 points (solid line) and padded to 128 points using the value f_{114}. The discrete wavelet and Fourier transforms are applied, and the resulting coefficients are truncated to 16 terms. The dash–dot and dotted lines shows the function after application of the inverse discrete wavelet and Fourier transforms, respectively.

$$\mathbf{C}_k = \mathbf{C}_{i,j}, \tag{20}$$

where $k = i + N_x(j - 1)$, and $N_x = 72$ is the number of grid points in the zonal direction. Thus, the first 72 points on either axis represent the error covariances at the most southerly grid points. The main diagonal represents the error variance as usual, and the last 72 points represent the most northerly grid points. The covariance between single point and the rest of the domain makes up a single column in the covariance, as shown in Fig. 3. The next column in the matrix will be the covariance for the grid point just to the right, and so on.

The transform from the physical space covariance to wavelet representation is then

$$\hat{\mathbf{P}} = \mathbf{WPW}^{\mathrm{T}}, \tag{21}$$

where $\mathbf{P}$ and $\hat{\mathbf{P}}$ are the physical and wavelet space covariance matrices, and $\mathbf{W}$, $\mathbf{W}^{\mathrm{T}}$ represent the wavelet transform and its transpose. For an $N \times N$ matrix, the transform requires $O(N^2 \log N)$ operations. In the present case the total number of grid points is 3312, but the matrix will be padded to obtain a power of 2 (as is needed by the discrete wavelet transform), so we will have $N = 4096$ after padding.

The structure of $\mathbf{P}^f$ has significant implications for truncation of $\hat{\mathbf{P}}^f$. Since the zonal direction is the inner loop in (20), the finest scales in $\mathbf{P}^f$ represent the smallest-scale structures in the x direction, while y-direction variations are at larger scales. Therefore a truncation to $L = 1024$ coefficients will mean that the two finest scales ($j = 11, 12$, corresponding to 5°–10° resolution) are removed in the zonal direction. Because of the structure of the error covariance, no information in the meridional direction is lost by this truncation. Since the steepest gradients in both the constituent field and the error covariance are in the meridional direction, we expect the Kalman filter to be most sensitive to loss of spatial information in this direction.

4. Description of the assimilation scheme in wavelet space

We consider two schemes for carrying out the propagation of error covariances in wavelet space. In the first, the analysis covariance $\mathbf{P}^a$ and the transport model $\mathbf{M}$ are projected onto a wavelet basis and the covariance is propagated in wavelet space. The covariance is then transformed back to physical space, and the analysis covariance and Kalman gain are calculated. The second approach leaves the analysis covariance in wavelet space and also calculates the Kalman gain in wavelet space. The former ensures that the nonuniform observation locations helps to introduce new wavelet coefficients in adaptive schemes, while the latter eliminates the need for forward and backward wavelet transforms at each analysis time.

FIG. 3. Schematic representation of how the error covariance for (bottom) a single point becomes (top) a single column of the whole error covariance matrix.

We use a wavelet transform that comes from the family of compactly supported orthonormal wavelets introduced by Daubechies (1988). The fast discrete transform of Mallat (1989) is employed and applied to the error covariance,

$$\hat{\mathbf{P}}^a = \mathbf{W}\mathbf{P}^a\mathbf{W}^\mathsf{T} \tag{22}$$

where $\mathbf{W}$ and $\mathbf{W}^\mathsf{T}$ represent the wavelet transform and its transpose, respectively, and $\hat{\mathbf{P}}^a$ is the wavelet representation of $\mathbf{P}^a$.

Computation of the model dynamics in wavelet space requires a bit more work because it is an operator rather than a matrix. However, $\mathbf{M}_k$ can be computed by applying the model operator to the identity matrix since it results in a matrix that carries out the same operations when applied to a vector. We then compute the wavelet representation of the model dynamics

$$\hat{\mathbf{M}}_k = \mathbf{W}\mathbf{M}_k\mathbf{W}^\mathsf{T}. \tag{23}$$

The approximation of the covariance and model dynamics is made through truncation of (22) and (23) in order to make use of the efficient representation of localized structures afforded by the wavelet representation. We consider two approaches to wavelet representation of covariances: (a) forecast covariance in wavelet space/analysis covariance in physical space and (b) both forecast and analysis covariances in wavelet space. Experiments using these approaches will be described in the next section.

a. Forecast in wavelet space/analysis in physical space

Updating of $\boldsymbol{\mu}^f$ remains the same,

$$\boldsymbol{\mu}^f_{k+1} = \mathbf{M}_k\boldsymbol{\mu}^a_k, \tag{24}$$

while the previous analysis error covariance, model dy-

namics, and model error are projected onto the wavelet basis,

$$\hat{\mathbf{P}}^a = \mathbf{W}\mathbf{P}_k^a\mathbf{W}^T, \tag{25}$$

$$\hat{\mathbf{M}}_k = \mathbf{W}\mathbf{M}_k\mathbf{W}^T, \tag{26}$$

$$\hat{\mathbf{Q}}_k = \mathbf{W}\mathbf{Q}_k\mathbf{W}^T. \tag{27}$$

The covariances and model dynamics are then truncated to L equations, based on either scale or error variance criteria. The truncated covariance is then updated,

$$(\hat{\mathbf{P}}_{k+1}^f)_L = (\hat{\mathbf{M}}_k)_L(\hat{\mathbf{P}}_k^a)_L(\hat{\mathbf{M}}_k^T)_L + (\hat{\mathbf{Q}}_k)_L. \tag{28}$$

The forecast error covariance is then transformed back to physical space,

$$\mathbf{P}_{k+1}^f = \mathbf{W}^T\hat{\mathbf{P}}_{k+1}^f\mathbf{W}, \tag{29}$$

and the analysis step

$$\boldsymbol{\mu}_{k+1}^a = \boldsymbol{\mu}_{k+1}^f + \mathbf{K}_{k+1}(\boldsymbol{\mu}_{k+1}^o - \mathbf{H}_{k+1}\boldsymbol{\mu}_{k+1}^f), \tag{30}$$

$$\mathbf{P}_{k+1}^a = (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1})[(\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1})\mathbf{P}_{k+1}^f]^T$$
$$+ \mathbf{K}_{k+1}\mathbf{R}_{k+1}^o\mathbf{K}_{k+1}^T, \tag{31}$$

where the Kalman gain matrix $\mathbf{K}_k$ is given by

$$\mathbf{K}_{k+1} = (\mathbf{H}_{k+1}\mathbf{P}_{k+1}^f)^T(\mathbf{H}_{k+1}\mathbf{P}_{k+1}^f\mathbf{H}_{k+1}^T + \mathbf{R}_{k+1}^o)^{-1}. \tag{32}$$

b. Forecast and analysis covariance updated in wavelet space

This scheme eliminates the need to transform the covariances between physical and wavelet space each analysis time, but what about $\mathbf{Q}_k$ and $\mathbf{M}_k$? Because $\mathbf{M}_k$ depends on the current winds and is continually changing, Eq. (26) is still needed, while $\mathbf{Q}_k$ is calculated by Eq. (4) as the product of the original correlation and current variance field. A computationally more efficient scheme for calculating $\hat{\mathbf{Q}}_k$ can be formulated by noting that

$$\mathbf{W}\mathbf{Q}_k\mathbf{W}^T = \delta^2(\mathbf{W}[\boldsymbol{\mu}_k]\mathbf{W}^T)(\mathbf{W}\mathbf{C}^q\mathbf{W}^T)(\mathbf{W}[\boldsymbol{\mu}_k]\mathbf{W}^T), \tag{33}$$

where $[\boldsymbol{\mu}_k]$ is a matrix with diagonal elements μ_k. The term $\mathbf{W}\mathbf{C}^q\mathbf{W}^T$ need only be computed once while the first and third terms of (33) must be recomputed each analysis time.

We also note that, since only L rows and columns of each matrix are being retained, the computation of the forward wavelet transforms can be significantly reduced by making use of the structure of the Mallat tree algorithm when the truncation criterion is by scale. The tree algorithm is a series of "averages" and "differences" taken at the finest scales first. This average on each scale is used to calculate the average and difference on the next coarser scale, and the complete set of differences makes up the set of wavelet coefficients. If we are retaining only the L largest coefficients, then only averages need to be taken for the first $N - L$ coefficients.

Since $L \ll N$ for our assimilation, the difference calculation is done for only a small fraction of the matrix.

The assimilation scheme again starts with an update of $\boldsymbol{\mu}^f$,

$$\boldsymbol{\mu}_{k+1}^f = \mathbf{M}_k\boldsymbol{\mu}_k^a. \tag{34}$$

Project the model dynamics onto the wavelet basis,

$$\hat{\mathbf{M}}_k = \mathbf{W}\mathbf{M}_k\mathbf{W}^T. \tag{35}$$

Project the mixing ratio contribution onto the wavelet basis and compute new model error in wavelet space,

$$\hat{\mathbf{Q}}_k = \delta^2(\mathbf{W}[\boldsymbol{\mu}_k]\mathbf{W}^T)\hat{\mathbf{C}}^q(\mathbf{W}[\boldsymbol{\mu}_k]\mathbf{W}^T), \tag{36}$$

where $\hat{\mathbf{C}}^q$ is the initial error correlation model in wavelet space.

Truncate dynamics and model error to L terms and calculate new forecast error covariance in wavelet space,

$$(\hat{\mathbf{P}}_{k+1}^f)_L = (\hat{\mathbf{M}}_k)_L(\mathbf{P}_k^a)_L(\hat{\mathbf{M}}_k^T)_L + (\hat{\mathbf{Q}}_k)_L. \tag{37}$$

Calculate analysis error covariance in wavelet space:

$$\hat{\mathbf{P}}_{k+1}^a = (\mathbf{I} - \hat{\mathbf{K}}_{k+1}\hat{\mathbf{H}}_{k+1})[(\mathbf{I} - \hat{\mathbf{K}}_{k+1}\hat{\mathbf{H}}_{k+1})\hat{\mathbf{P}}_{k+1}^f]^T$$
$$+ \hat{\mathbf{K}}_{k+1}\mathbf{R}_{k+1}^o\hat{\mathbf{K}}_k^T, \tag{38}$$

$$\hat{\mathbf{K}}_{k+1} = \mathbf{W}\mathbf{K}_{k+1}, \quad \text{and} \tag{39}$$

$$\hat{\mathbf{H}}_{k+1} = \mathbf{H}_{k+1}\mathbf{W}^T \tag{40}$$

are the Kalman gain and observation operator mixed physical/wavelet space. Notice that $\mathbf{R}_{k+1}^o$ is in physical space, because $\hat{\mathbf{K}}$ and $\hat{\mathbf{K}}^T$ are only right and left multiplied by the wavelet transform matrix, respectively.

5. Assimilation results

Success of any approximate scheme for evolving error covariances is ultimately measured by the accuracy of the analyses themselves and the reduction in computational cost that the scheme achieves. We also wish to understand how well the wavelets represent the covariance and dynamics matrices, and how the accuracy of this representation affects the assimilation accuracy. It is particularly important to understand the relationship between the loss of error covariance information and the decrease in assimilation accuracy that occurs as the number of coefficients retained is reduced.

In all cases, the computational grid is 72×46, so that each correlation matrix is 72×46 and the full covariance matrix is then $n \times n = 3312 \times 3312$. The fast wavelet transform requires that the system dimension be a power of 2, so we pad the covariance to $n \times n = 4096 \times 4096$. Truncation is carried out on the finest scales first. The finest scale is always half of the total number of coefficients; removing 2048 wavelet coefficients is equivalent to removing just the $j = 11$ scale in the zonal direction, for example.
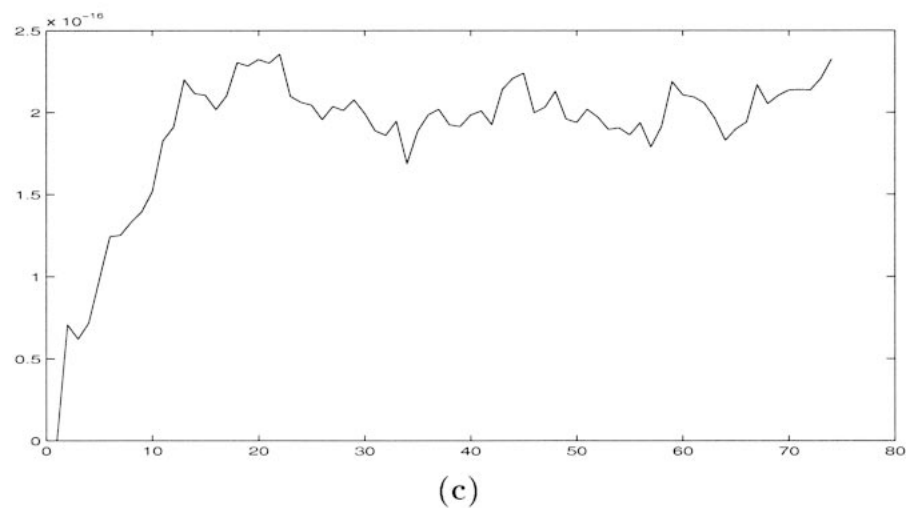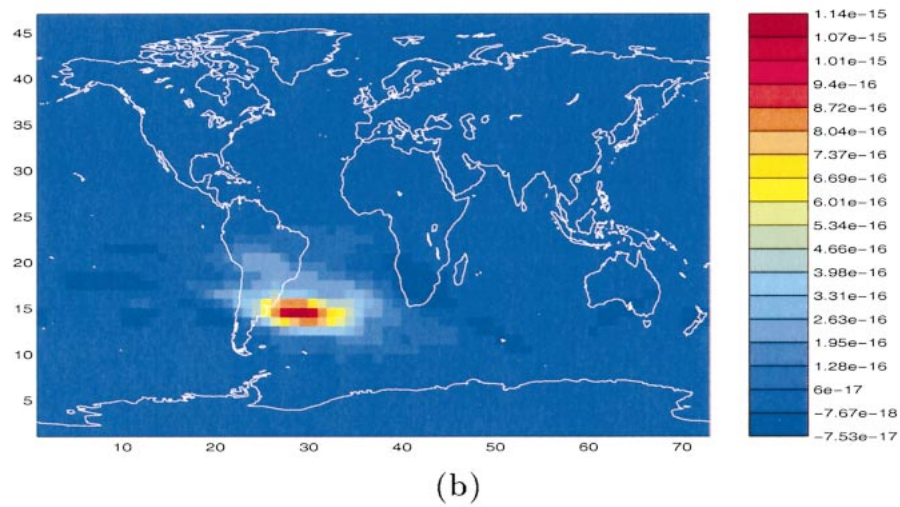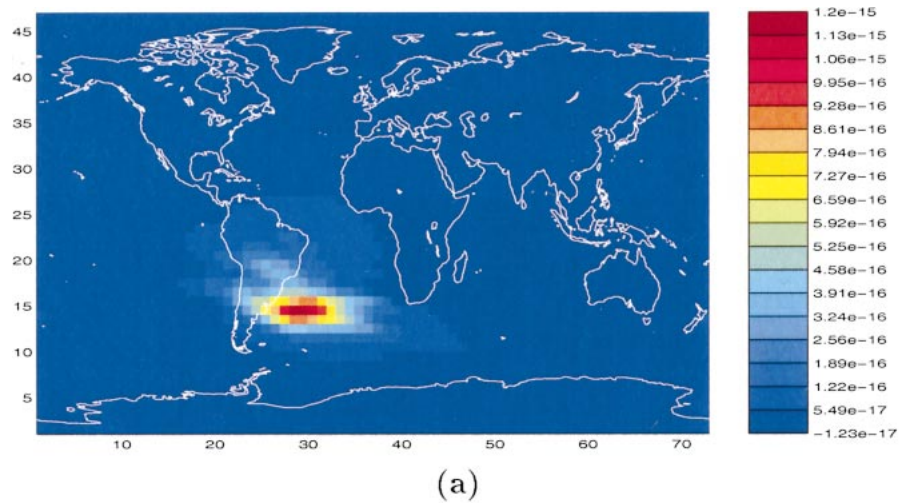
FIG. 4. Assimilation results with forecast error covariance updated in wavelet space and analysis error covariance calculated in physical space. (a) The error correlation after 8 days for the full system, and (b) the correlation after 8 days with a truncation of 1024 × 1024 wavelet coefficients. (c) The mean quadratic difference in the analysis constituent fields for the truncated and full systems. The x axis is number of 3-h analysis cycles. The total time is about 8 days.
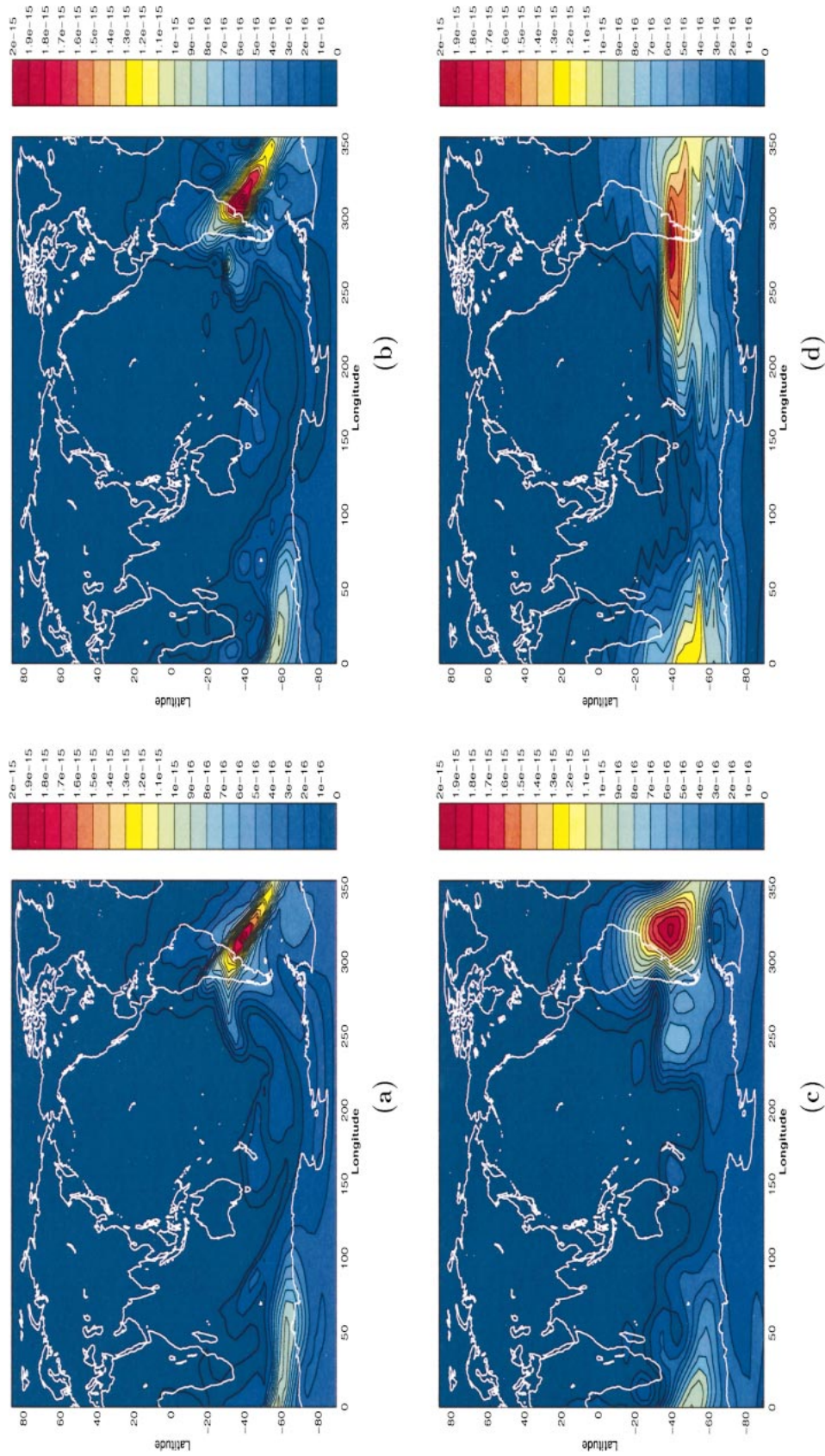
FIG. 5. (a) Full covariance propagated in physical space, with (b) 1024 × 1024 truncation, (c) 512 × 512 truncation, (d) 256 × 256 truncation. Error covariance for a point (43°S, 45°W) after 19-days assimilation with both analysis and forecast error covariances calculated in wavelet space is shown in (b)–(d).
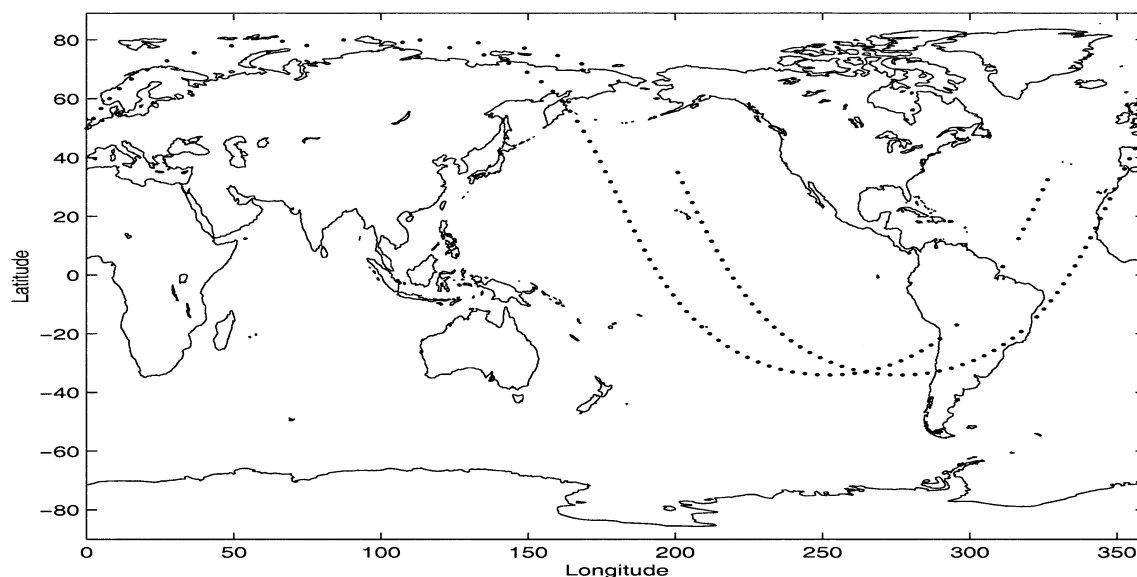
FIG. 6. Observation locations for the analysis time of the covariance shown in Fig. 5, on day 19 of the assimilation.

a. Forecast covariance in wavelet space

Constituent assimilation using both the full and truncated error covariances has been carried out over a period of 19 days. Truncation to $L = 1024$, resulting in the removal of the two finest zonal scales ($j = 10,11$), was used. The correlation for a single point after 9 days of assimilation is shown in Fig. 4 for (a) the full system and (b) $L = 1024$. The correlation that results from the truncated system is seen to be qualitatively and quantitatively very close to the full system correlation. The rms difference in the constituent fields between the truncated and full systems is shown in Fig. 4c. The error incurred by truncation of the covariance and the TLM shows an initial rapid increase that levels off around $2\frac{1}{2}$ days. This stabilization indicates that the approximation scheme will not result in long-term error growth relative to the full Kalman filter system.

Computational savings for this scheme, however, were found to be rather small. While truncation of the covariance propagation equation reduces the operation count by $(L/N)^2 = (1024/3312)^2 = 1/10$, the scheme requires a forward and back wavelet transform on the entire covariance and a forward transform on the propagator each analysis time. Each transform consists of $O(N_p^2)$ operations ($N_p = 4096$ is the padded vector length), which in this case is carried out twice every 12 time steps. The computational savings over one analysis cycle due to the wavelet truncation is nearly equal to the computational cost of the wavelet transforms. Thus there is no significant savings if the analysis error covariance is calculated in physical space. A further reduction in L will only slightly decrease the computational cost of the propagation step relative to the cost of the wavelet transforms. We conclude that the analysis error covariance needs to be determined in wavelet space in order to realize any significant computational savings.

b. Forecast and analysis covariance in wavelet space

Constituent assimilation was carried out for 24 days (the same initial state as in the previous section) using the approximation scheme in (34)–(40), with $L = 1024$, 512, and 256. These are again compared with the benchmark assimilation that evolves the full error covariance in physical space. The impact of the misspecification of error covariance on the analysis field should be greatest near observation locations during a given analysis cycle. That is, if there are no observations in some region, no assimilation is carried out there and no differences in the analyzed field would be generated for the different approximations of the error covariance. We examine this impact by comparing full and approximated forecast covariances for the point (43°S, 45°W) at three different assimilation times. These can be divided into three different cases with respect to accuracy of error covariance approximation and the proximity of satellite observations. The first case occurs when there are observations in the vicinity of (43°S, 45°W) and the truncated wavelet expansions do not accurately approximate the error covariance. In the second case there are also nearby observations, but the covariance approximation is considerably better. The third case occurs when there are no observations nearby. We examine the error covariance between the node at (43°S, 45°W) and the rest of the domain on three different days (19, 21, and 22) of the assimilation and describe the relationship between the accuracy of local error covariance and the relative accuracy of the assimilated constituent field.

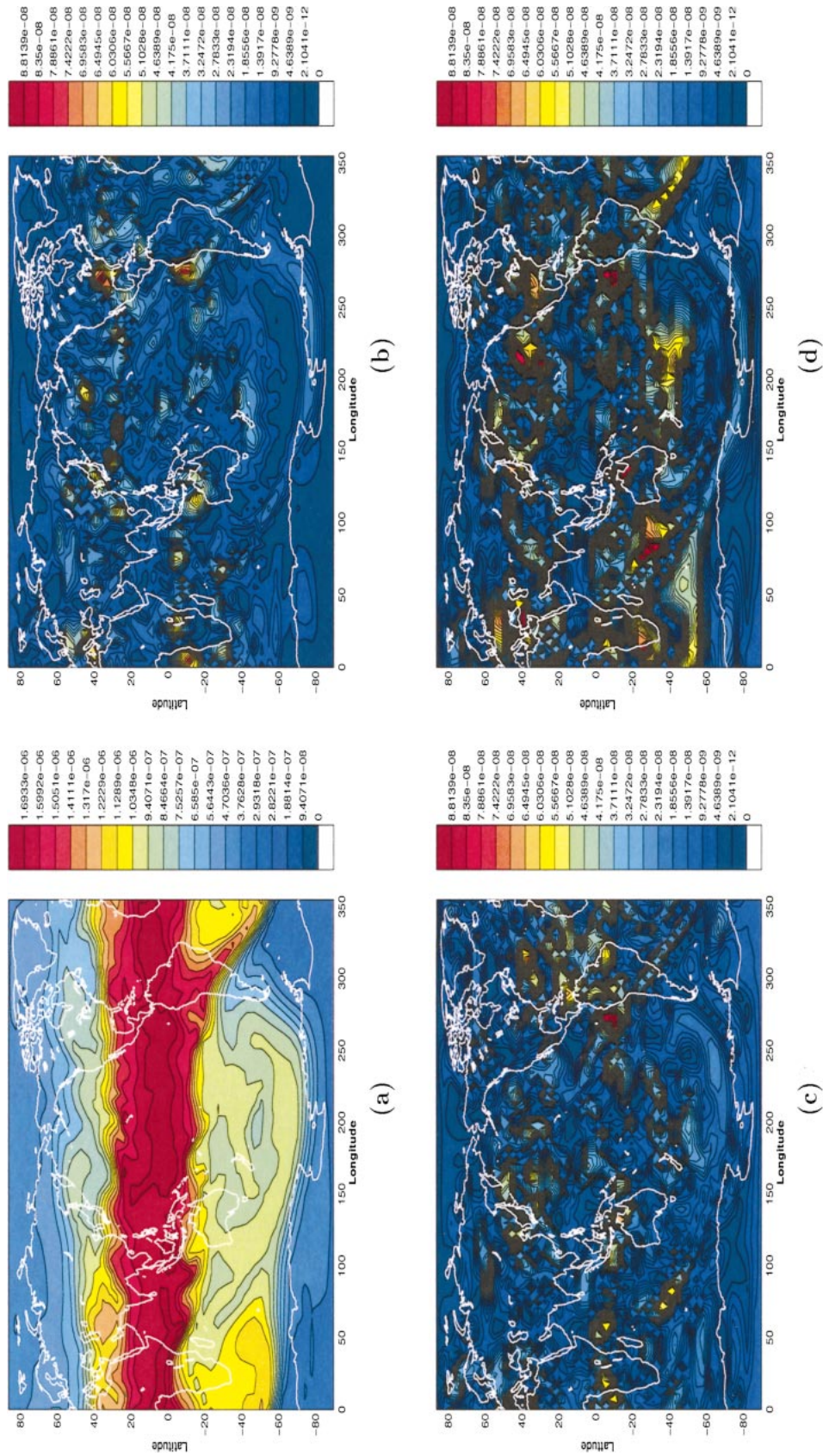A case 1 example can be seen in Fig. 5, which shows

FIG. 7. (a) Analysis constituent field at time step 1873 (1200 UTC on 19th day) with full error covariance propagation. Differences with respect to (a) are shown for wavelet coefficient truncations of (b) 1024 × 1024, (c) 512 × 512, and (d) 256 × 256.
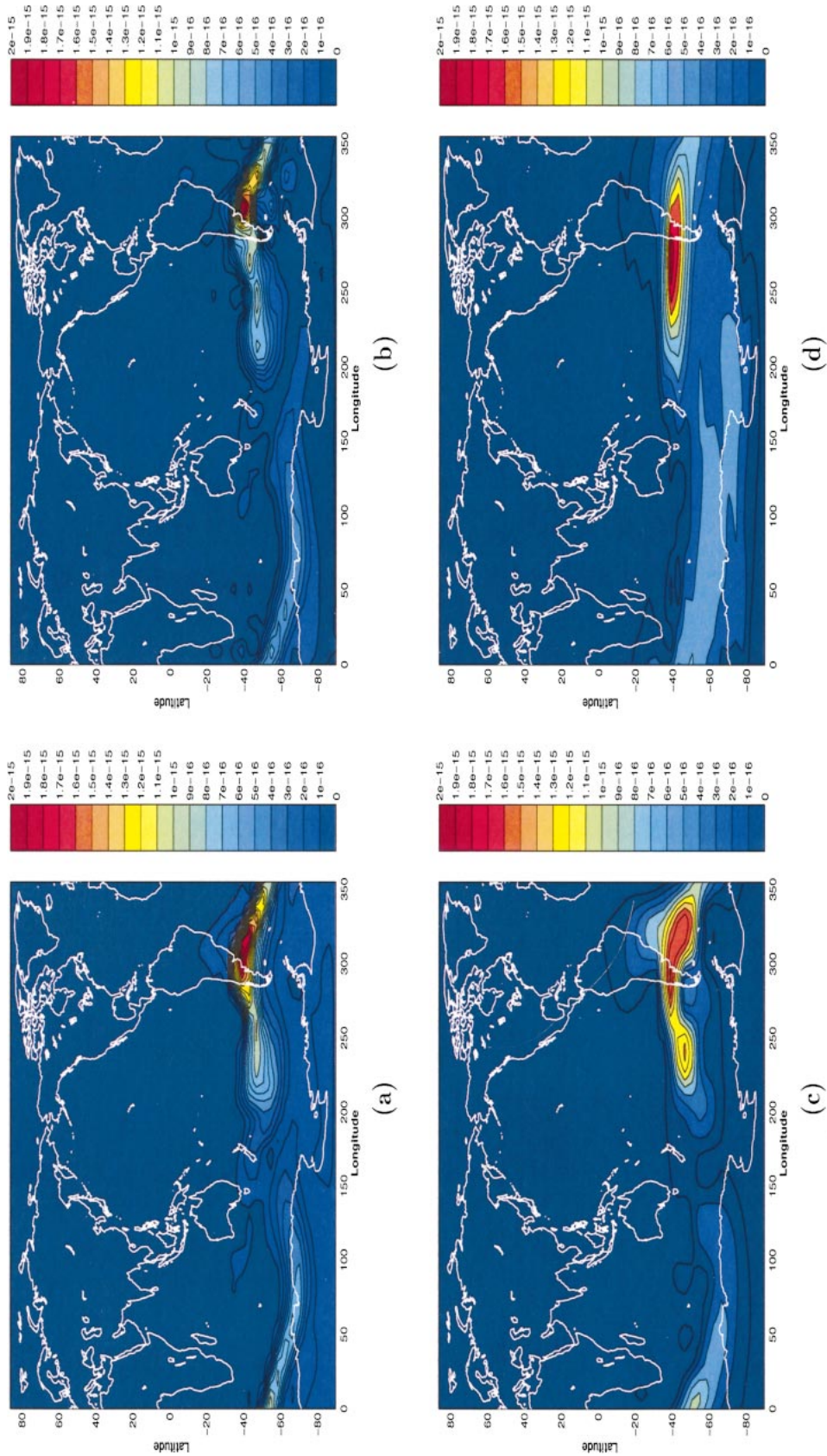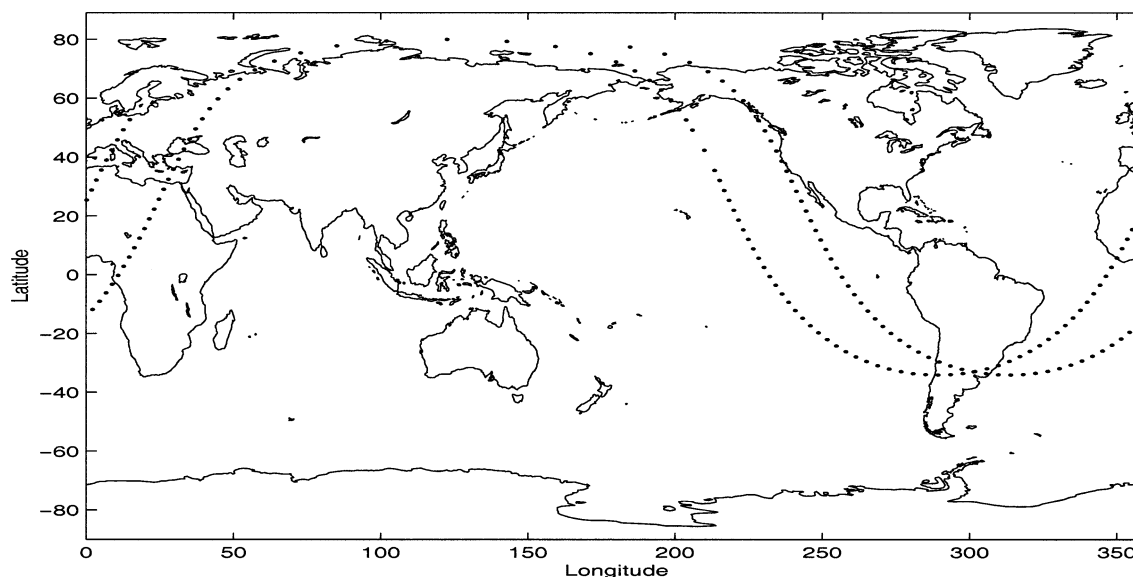
FIG. 8. Analysis error covariance for the point (43°S, 45°W) after 20-days assimilation with both analysis and forecast error covariances calculated in wavelet space. (a) Full covariance propagated, (b) 1024 × 1024 truncation, (c) 512 × 512 truncation, and (d) 256 × 256 truncation.

FIG. 9. Observation locations for the analysis time of the covariance shown in Fig. 8, on day 21 of the assimilation.

the forecast error covariance on the 19th day (time step 1873). Shown in Fig. 5 is (a) full error covariance evolution, along with wavelet space approximations that retain (b) 1024×1024, (c) 512×512, and (d) 256×256 coefficients. The important structure of the covariance for this point is the stretching to the southeast that extends the 50% correlation contour as far as the 0° longitude and the 33% contour to around 50°W. The 1024×1024 coefficient approximation accurately reproduces this structure, but both the 512×512 and 256×256 approximations are unable to represent the general shape. Figure 6 shows observation locations for the analysis carried out using the forecast error covariances in Fig. 5. Several observations lie within the region where the correlations are above 50% for full covariance evolution. In particular, several of the observations over South America should impact the analysis along the length of the narrow region from (35°S, 60°W) to (60°S, 0). Figure 7 shows the CH_4 concentration analysis field on day 19 corresponding to the covariance fields in Fig. 5. Figure 7a shows the analysis field that results from the evolution of the full error covariance, while Figs. 7b–d show the difference between the analysis fields that result from the evolution of truncated (1024×1024, 512×512, and 256×256) wavelet expansions for the error covariances. (It is important to recall that in all cases the constituent field is evolved in physical space and does not involve any truncation.) The finger of high concentration extending from South America into the South Atlantic in Fig. 7a corresponds to the covariance region in Fig. 5. Figure 7b shows only small differences in the constituent field (in the region of interest) for the 1024×1024 error covariance truncation, while Figs. 7c and 7d show significantly larger differences due to the relatively poor covariance representation. Note that outside the region of interest, there may

be little difference in the errors depending on the local success of representing the error covariances.

An example of case 2 (good covariance approximation and with observations nearby) can be seen in Fig. 8. Now all of the truncated wavelet representations of the covariance make reasonable approximations to both the shape and magnitude. As a result, the nearby observations (Fig. 9) are assimilated to give constituent analyses (Fig. 10) that show significantly smaller differences with the full system in the region near the covariance plot of Fig. 5. Case 3 occurs in the same location on the 22d day of the assimilation, with error covariances shown in Fig. 11 and a satellite observation track (Fig. 12, which does not cross this covariance region). The constituent analyses (Fig. 13) again show little difference in the region of interest.

On day 19 of the assimilation, the local error covariance has significant gradients in the zonal direction. This information is lost in the most severe truncations, leading to the differences found in the constituent field at that time. Later (day 21), the covariance in the same region is primarily stretched in the zonal direction and even the most severe truncation leads to a good approximation of the covariance field.

The relationship between zonal gradients in the constituent field and assimilation accuracy are difficult to show because even when there is no zonal gradient in some region, earlier errors that have been advected (by winds) to that region will show as local error. A somewhat more definite relationship can be made between zonal constituent gradients and misrepresentation of the error covariance by the reduced rank Kalman filter. Figure 14 shows a scatterplot of the normalized zonal concentration gradient, $(dC/dx)/C$, versus the rms difference between full and 512×512 error covariances over a 15-day period. For any level of error in the covariance
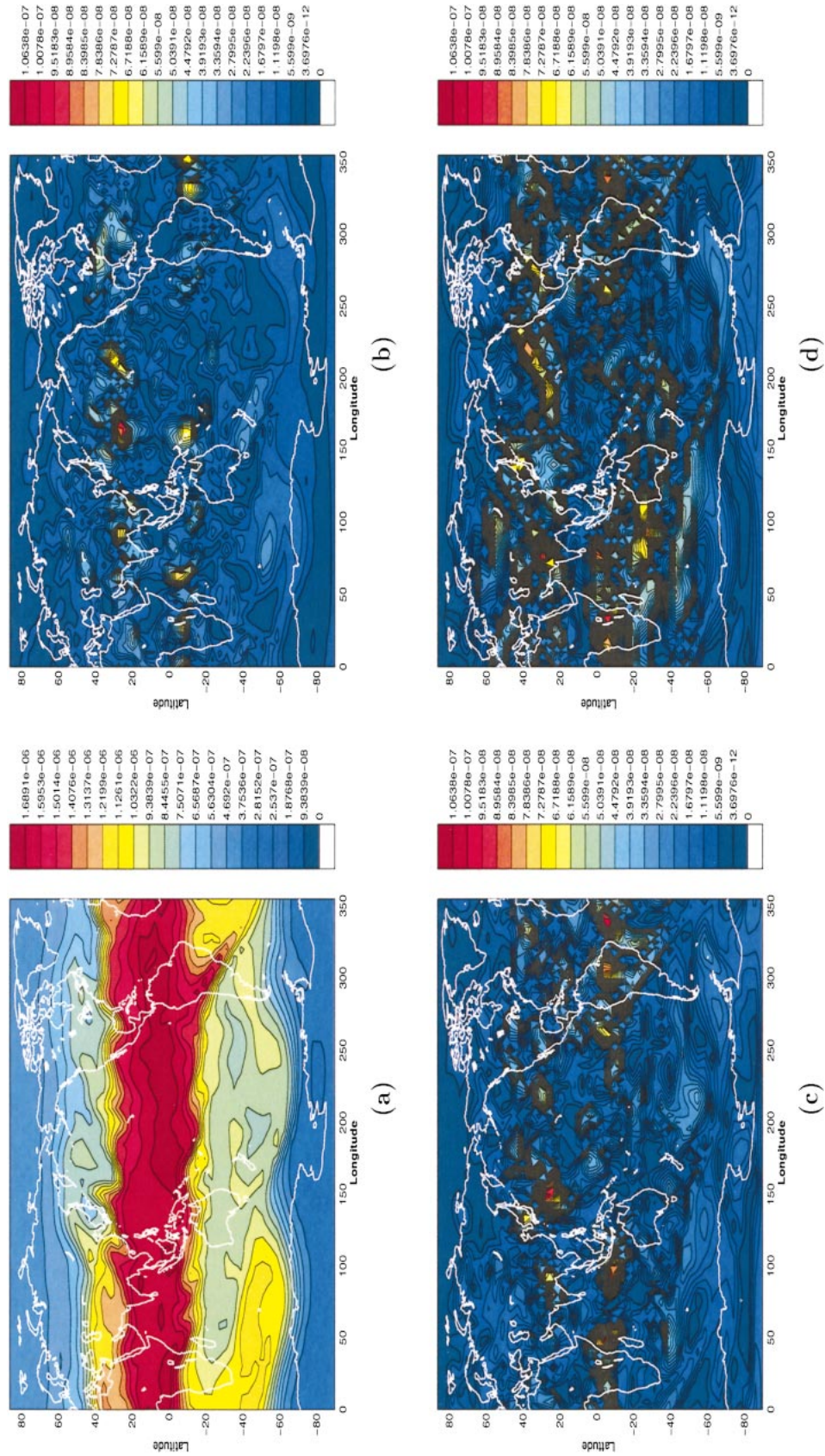
FIG. 10. (a) Analysis constituent field at time step 2053 (1200 UTC on the 21st day) with full error covariance propagation. Errors with respect to (a) are shown for wavelet coefficient truncations of (b) 1024 × 1024, (c) 512 × 512, and (d) 256 × 256.
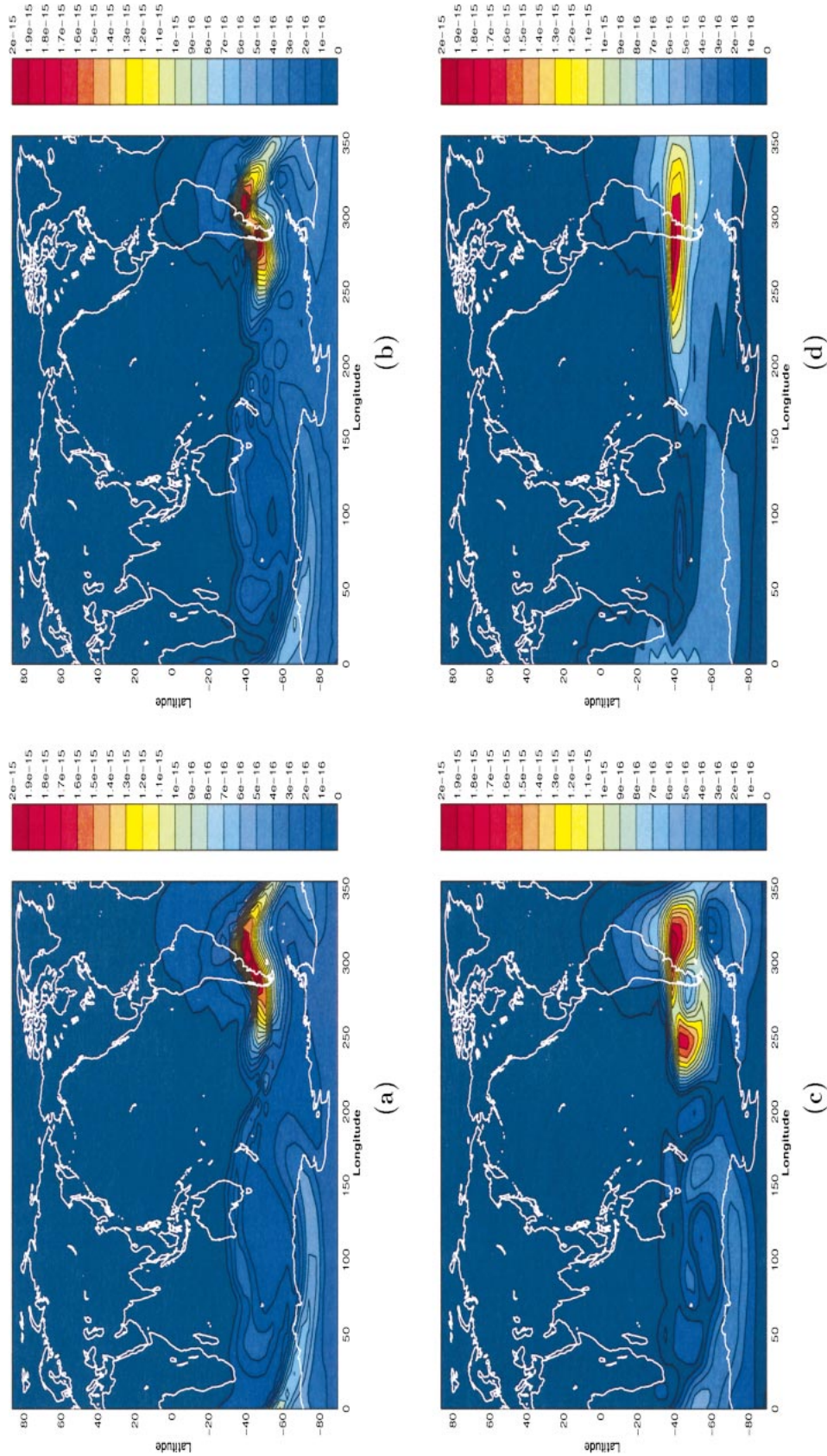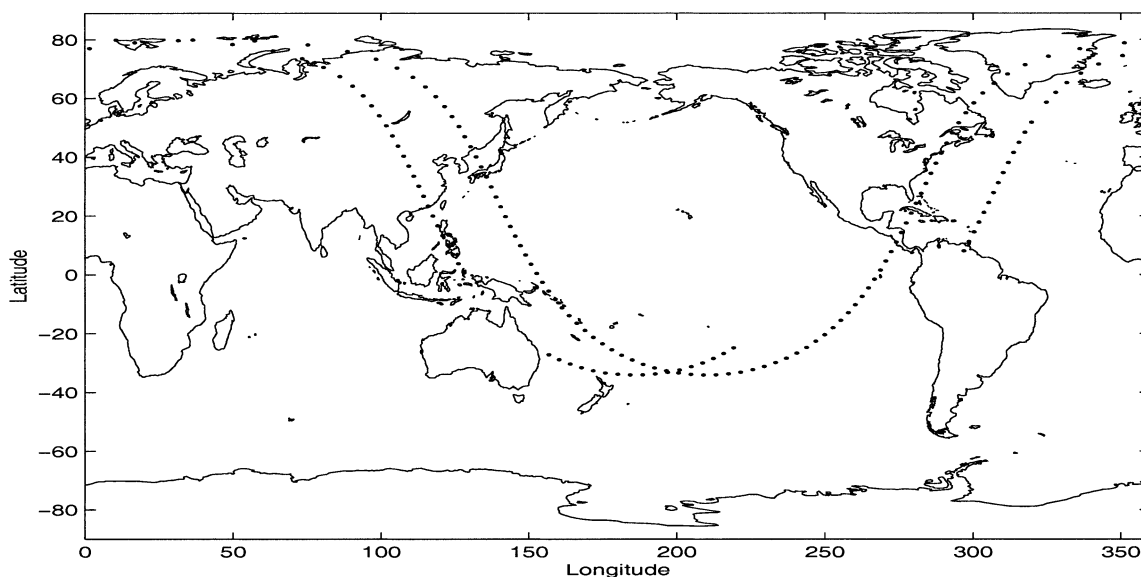
FIG. 11. Error covariance for the point (43°S, 45°W) after 22-days assimilation with both analysis and forecast error covariances calculated in wavelet space. (a) Full covariance propagated, (b) 1024 × 1024 truncation, (c) 512 × 512 truncation, and (d) 256 × 256 truncation.

FIG. 12. Observation locations for the time of the error covariance shown in Fig. 11.

matrix, there is a maximum value for $(dC/dx)/C$ that increases from about 3×10^{-14} to 10^{-12} km^{-1} as the rms difference increases from 10^{-15} to 2×10^{-14} km^{-1}. This indicates a general correlation between the minimum error and the zonal constituent gradient. The overall scatter in the figure is most likely due to advected errors that act to increase the rms error in the covariance matrix in some locations even where the zonal gradient is small.

Furthermore, not all of the differences in the constituent field can be ascribed to loss of resolution in the zonal direction. Farrell and Ioannou (2001) discussed the issue of dynamics truncation, which can lead to errors in covariance propagation at later times because some growing modes may be lost. Also, truncation errors in zonal direction error covariance can lead to more general errors through the propagator, **M**.

It is also useful to plot the time evolution of the mean absolute error,

$$E_{\text{mean}} = \frac{1}{A_{\text{tot}}} \sum_{i,j} |\mu i, j - \mu_{\text{ref}}(i, j)| \Delta A_{i,j}, \qquad (41)$$

in the constituent field with respect to the full Kalman filter system, shown in Fig. 15. We define $\Delta A_{i,j}$ as the area increment associated with node (i, j). The mean error in the $L = 1024$ case is seen to stop increasing and actual decrease starting around the fifth day, while the errors $L = 512$ and $L = 256$ cases show a decrease followed by an increase starting at day 17. This last feature is a result of a loss of data starting on day 12 and lasting through day 15. Because there are no observations during this interval, there are no analysis increments, which would differ depending on the accuracy of the evolved error covariance (thus reinforcing the differences in the constituent fields). Once the ob-

servations resume, the analysis increments resume, creating the difference between the benchmark and the approximate covariance propagation. However, there is no overall growth in errors between day 5 and day 24.

The reductions in computational costs associated with this scheme over the course of a 24-day assimilation are about 80% for $L = 1024$, 93% for $L = 512$, and 96% for $L = 256$. Each of these are compared with the Kalman filter system with error covariance evolved in physical space. The computational cost is roughly proportional to L^2, in spite of the additional expense of constructing the propagator in matrix form and projecting it onto the wavelet basis each analysis time.

6. Discussion and conclusions

We have implemented a Kalman filter system for the assimilation of chemical constituent observations from limb-sounding instruments in which the error covariances are propagated forward in terms of their wavelet coefficients. This numerical scheme has been compared with the benchmark system of Ménard et al. (2000) with several different degrees of truncation of the wavelet coefficients. The truncation removes the finest scales first, and only in the zonal direction. Thus, as the number of retained coefficients in the covariance propagation are reduced, increasingly coarser scales are eliminated. The important information that can be obtained from these experiments is the relationship between changes in the error covariances and the constituent field itself. Since error variances are larger in regions of steep gradients in the constituent field, the two-point covariances will tend to be aligned in the zonal direction, as is seen in Fig. 8. Structures with this shape will require a lower
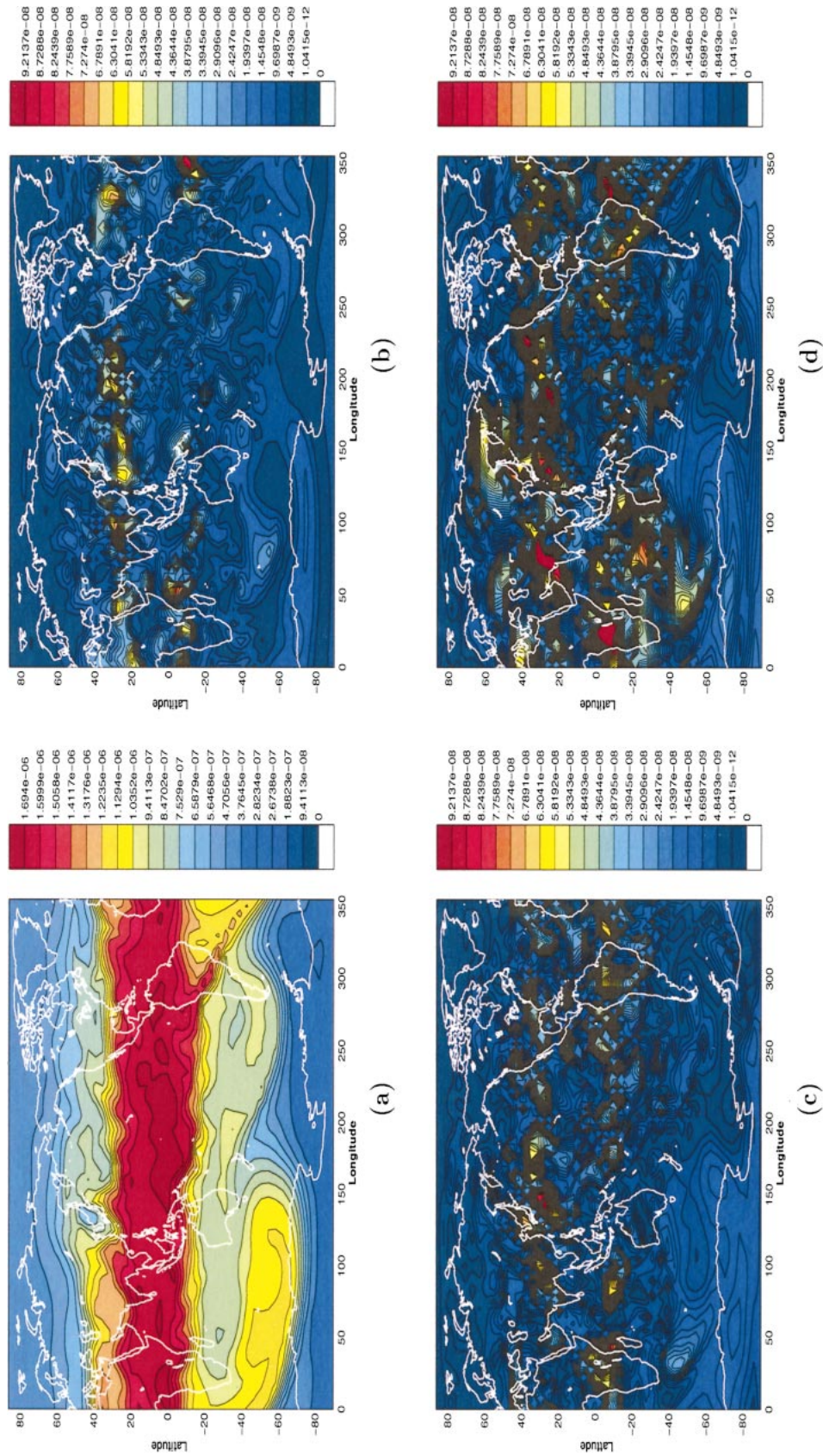
FIG. 13. (a) Analysis constituent field at time step 2173 (1200 UTC on the 22d day) with full error covariance propagation. Errors with respect to (a) are shown for wavelet coefficient truncations of (b) 1024 × 1024, (c) 512 × 512, and (d) 256 × 256.
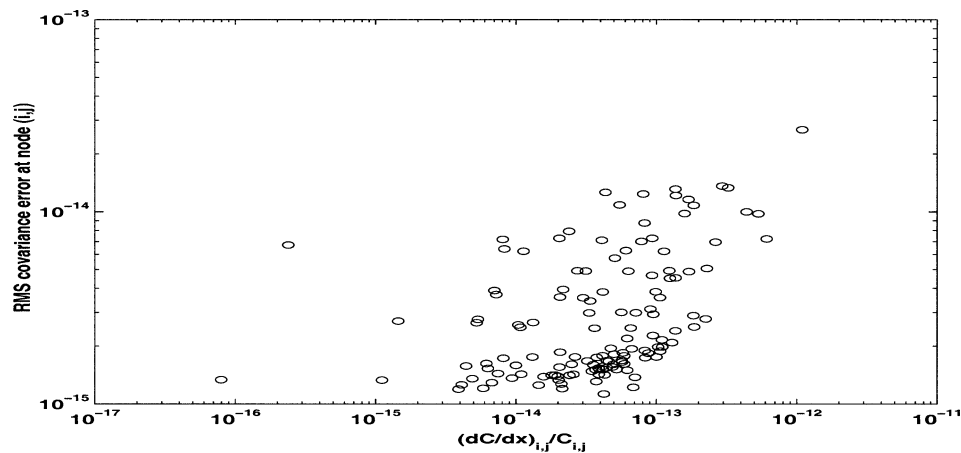
FIG. 14. Log-scale scatterplot of dC/dx vs rms difference between full error covariance evolution and the 512×512 approximation for the grid point located at (43°S, 45°W), over 15 days of the assimilation. Each circle represents a single assimilation time.

resolution in the zonal direction, thus allowing the finest-scale coefficients in this direction to be eliminated.

These experiments show that as the resolution of the error covariance propagation is reduced, errors in the constituent field will increase more where the gradient in the constituent field is largest. The largest representation errors in the error covariances are found to occur where the gradients are steepest in the zonal direction. However, transport of errors in the constituent field complicates the relationship between zonal gradients and assimilation errors.

While this work does not consider adaptive schemes that vary the local scale retained according to a priori estimates of the truncation error, these assimilation results give some insight as to how such a scheme might work. We can take advantage of the fact that steep gra-

dients in the constituent field give rise to larger error variances as well as sharp features in the covariance field. The constituent forecast is updated before the error covariance, so it is available for estimating the location of small-scale features in the covariance field. An adaptive scheme of this could be combined with an error covariance estimate from either a flow-dependent scheme (Riishøjgaard 1998) or a low-order ensemble Kalman filter (e.g., Houtekamer and Mitchell 1998). This would allow the number of wavelet coefficients to be further reduced with a very modest computational expense. We have also not considered alternate wavelet expansions that might better fit a spherical geometry (Göttelmann 1999). However, we have not found any significant errors generated by the pole problem, most likely because there is little in the way of observations
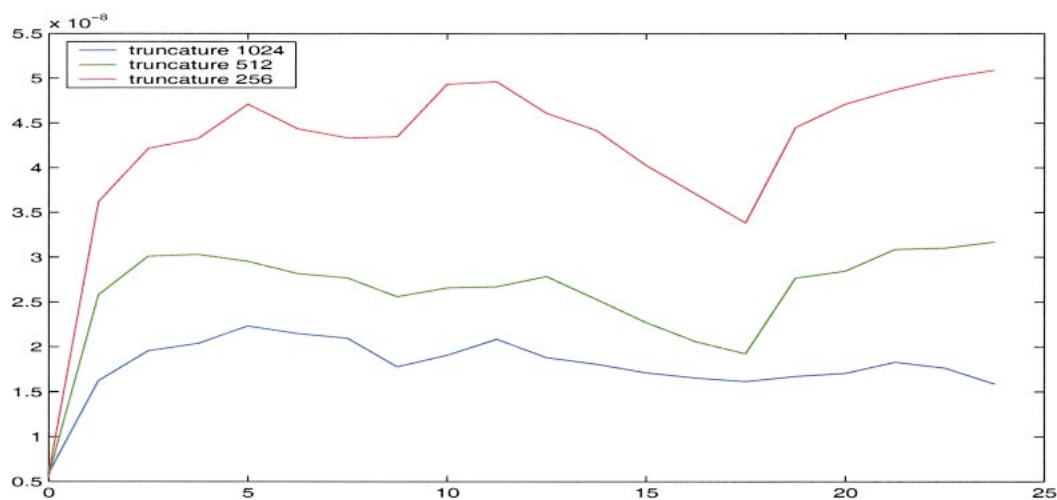

FIG. 15. Mean absolute error for the constituent field relative to the benchmark (full covariance evolution) for a 24-day assimilation. The error is defined as $\epsilon = 1/A \; \Sigma_{i,j} \; |\mu(i, j) - \mu_{\mathrm{ref}}(i, j)| \, dA_{i,j}$, where $dA_{i,j}$ is the area associated with the (i, j)th grid point and A is the global area. Both the forecast and analysis covariances are calculated in wavelet space.

or gradients in the constituent field near the poles. These issues will need to be addressed in the development of more general assimilation systems.

REFERENCES

Cane, M. A., A. Kaplan, R. N. Miller, B. Tang, E. C. Hackert, and A. J. Busaslacchi, 1996: Mapping tropical Pacific sea level: Data assimilation via a reduced state space Kalman filter. *J. Geophys. Res.,* **101C,** 22 599–22 617.

Chin, T. M., A. J. Mariano, and E. P. Chassignet, 1999: Spatial regression and multiscale approximations for sequential data assimilation in ocean models. *J. Geophys. Res.,* **104,** 7991–8014.

Cohn, S. E., and R. Todling, 1996: Approximate data assimilation schemes for stable and unstable dynamics. *J. Meteor. Soc. Japan,* **74,** 63–75.

Daubechies, I., 1988: Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.,* **41,** 909–996.

Ehrendorfer, M., and J. Tribbia, 1997: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.,* **54,** 286–313.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99,** 10 143–10 162.

Farge, M., 1992: Wavelet transforms and their application to turbulence. *Annu. Rev. Fluid Mech.,* **24,** 395–457.

Farrell, B. F., and P. J. Ioannou, 2001: Accurate low-dimensional approximation of the linear dynamics of fluid flow. *J. Atmos. Sci.,* **58,** 2771–2789.

Fukumori, I., and P. Malanotte-Rizzoli, 1995: An approximate Kalman filter for ocean data assimilation: An example with an idealized Gulf Stream model. *J. Geophys. Res.,* **100C,** 6777–6794.

Göttelmann, J., 1999: Locally supported wavelets on manifolds with applications to the 2D sphere. *Appl. Comput. Harmon. Anal.,* **7,** 1–33.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

Lin, S.-J., and R. B. Rood, 1996: Multidimensional flux-form semi-Lagrangian transport schemes. *Mon. Wea. Rev.,* **124,** 2046–2070.

Mallat, S. G., 1989: A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.,* **11,** 675–693.

Ménard, R., and L.-P. Chang, 2000: Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: χ^2-validated results and analysis of variance and correlation dynamics. *Mon. Wea. Rev.,* **128,** 2672–2686.

——, S. E. Cohn, L.-P. Chang, and P. M. Lyster, 2000: Stratospheric assimilation of chemical tracer observations using a Kalman filter. Part I: Formulation. *Mon. Wea. Rev.,* **128,** 2654–2671.

Riishøjgaard, L. P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus,* **50A,** 42–57.

Tangborn, A., and S. Q. Zhang, 2000: Wavelet transform adapted to an approximate Kalman filter system. *Appl. Numer. Math.,* **33,** 307–316.

Tippet, M. K., S. E. Cohn, R. Todling, and D. Marchesin, 2000: Low-dimensional representation of error covariance. *Tellus,* **52A,** 533–553.